*Malaysian Journal of Applied Sciences* 2018, **Vol 3(2)**: 21-33

eISSN 0127-9246 (Online)

UNIVERSITI SULTAN ZAINAL ABIDIN

**Malaysian Journal of Applied Sciences**

# Extended Cox Modelling of Survival Data with Guarantee Time

***K. A. Adeleke [a], A. A. Abiodun [b] and R. A. Ipinyomi [b]***

[a] Department of Mathematics, Obafemi Awolowo University, Ile-ife, Nigeria.
[b] Department of Statistics, University of Ilorin, Nigeria.

*Corresponding author: adedayobright@gmail.com

*Received: 08/05/2018, Accepted: 10/12/2018*

## Abstract

Proportional Hazard regression model for censored survival data often specifies that covariates have a proportional fixed effect on the hazard function of the lifetime distribution of a subject. A modification of the proportional hazards model of Cox (1972) to accommodate the non-proportional effect on hazard with a time-varying covariate and the introduction of guarantee time into the Weibull distributed baseline hazard function. Simulations were conducted to investigate properties of the models. Our approach had shown to have the best asymptotic properties in a simulation study with mean, Absolute Bias (AB) and Mean Square Error (MSE) of the parameter estimates for the models (under different levels of censoring and sample sizes) using simulated data.

**Keywords**: Cox model, Extended Cox Model, guarantee time, time varying covariate.

## Introduction

The assumption that covariates remain fixed over survival time may not always be true and current values of covariates may be more meaningful than its value when the study started Arasan and Lunn (2009). Failure in the assumption of proportionality leads to Non-proportional hazards models. This leads to the recent developments on the modelling of time-varying covariate and its effect. In most medical problems, proportional assumption failed because most of the factors are time-varying, thereby allowing applicability of NPH models. Examples of covariates that change in their values over time are blood pressure, exposure to radiation blood cell count. New extensions of the Cox model with time-varying covariate have been developed. There are a number of related articles on these models. The works of Leemis (1987, 1990) and Shihand and Leemis (1993) offered different frameworks on the Cox model with time-varying covariate following the accelerated life and proportional hazards models basically on a time-varying covariate. Zhou (2001) used an exponential distribution in conjunction with a transformation of the Cox model to including time-varying covariate. Failure time sometimes is modelled to include an initial threshold parameter (or guarantee parameter) before which it is assumed that failure cannot occur, Kalbfleisch and Prentice (2002).

The introduction of guarantee time into Exponential, Weibull and Gompertz distribution were considered by Lee and wang (2003) and lawless (2003). Scheike (2004), Wong et. al. (2006), Arasan and Lunn (2009) presented some developments that considered time-varying effect of covariates and also emphasized the use of semiparametric models where some effects are time-varying and some are time-fixed; thus giving the extended flexibility only for effects where a simple description is not possible. Occasionally, situations call for the inclusion of such a parameter. Bender et. al. (2005) generated survival times that follow Proportional hazards applied it to modelling characteristics of human mortality. Austin (2012) came up with the work on the generation of survival time that follows both proportional and non-proportional hazards model. This was an improvement over Bender et.al. (2005). He extended the work by considering non-proportional hazard and proportional hazard models i.e Semi-parametric hazards models. In this paper, we extend Cox proportional hazard models given by Cox (1972) as well as the work of Austin (2012) by incorporating a guarantee time as proposed by Lee and Wang (2003) using Weibull distributed baseline hazards. We performed simulation in order to investigate some asymptotic properties of the estimated parameters.

## Materials and Methods

Consider the guarantee time tg; tg>0, specified for a subject under study and defined as the time within which no death or failures can occur or the minimum survival time (Lee and Wang, 2003). A good real-life example of a scenario with guarantee time is the expected delivery time in the duration of pregnancy. Some pregnancies are delivered before the due date or a minimum number of weeks. Let ti= min (T; C), be the observed time to event (failure), where T is the follow-up time, with T <tg; T >tg; or T = tg and C is the censoring time which is often random. If tg= 0 then, the model reduces to proportional hazards model (Parametric or Cox). This study considers a situation where T<tg. We begin with an exponential function proposed by Lee and Wang (2003), where guarantee time is defined as the time within which no death/ birth/failures can occur or a minimum survival time. In their own case, the model reduces to the usual exponential function if tg= 0. We propose in this study a time-varying model with guarantee time. Let us first consider the usual exponential function as defined by Lee and Wang (2003) given by

$$f(x) = \begin{cases} \lambda \exp - \lambda(t), & t \geq 0 \\ 0, & t < 0 \end{cases} \tag{1}$$

Incorporating the guarantee time, we have:

$$f(x) = \begin{cases} \lambda \exp - \lambda(t - t_g), & t \geq t_g \\ 0, & t < t_g \end{cases} \tag{2}$$

with cummulative function given as

$$F(t) = \{1 - \exp - \lambda(t - t_g), \quad t \geq t_g$$

and the baseline hazard

$$h_o(t) = \lambda \tag{3}$$

Recall that the standard Cox (1972) model is given by:

$$h(t|x) = h_0(t)\exp(\beta' x) \tag{4}$$

and the corresponding Extended Cox model with time varying covariate z(t) is given by:

$$h(t|x) = h_0(t)\exp(\beta' x + \gamma' z(t)) \tag{5}$$

where $\beta$ and $\gamma = \beta_t$ are time fixed and time varying coefficients respectively. Clearly (3) is a component of (4) if the baseline survival function follows exponential distribution. To introduce guarantee time $t_g$ into the baseline hazard function of (5), we let $\tau = (t - t_g)$ ; for $t \geq t_g$ , we obtained:

$$h(t - t_g \,|\, z(x), x) = h_0(t - t_g)\exp[\beta' x + \gamma' z(t)]$$

$$h(\tau \,|\, z(t), x) = h_0(\tau)\exp[\beta' x + \gamma' z(t)] \qquad (6)$$

Equation (6) is the Extended Cox Model with guarantee time $t_g$ , which is also a member of semiparametric non-proportional hazard models. As noted by Bender et al (2005). and Austin (2012), the parameters required for each distribution, the hazard function, the cumulative hazard function, inverse cumulative hazard function and the formula for simulating survival times for each distribution in the setting of time -invariant covariates are described in Table 1. A parametrization of Bender et al (2005) is used for the Weibull distribution.

***Simulation Study***

Let $Y_i$ be a random variable with cumulative distribution function F and density function f, then U=F(Y) is uniform distributed random variable U~U(0,1). Moreover, if U~U(0,1), then, 1-U ~ U(0,1) also (Mood and Graybill (1974).Thus let T be the survival time of (4), then it follows that:

$$U=\exp[-H_0(T)\exp(B'x)]$$

~U(0,1), a uniform random variable and survival time T can be expressed as:

$$T = H^{-1}[-\log(U)\exp(-\beta' x)] \qquad (7)$$

See Bender et al. (2005). The survival time generated in their work follows a proportional hazard assumption as presented in (Table 1). The works of Zhou (2001), Austin(2012) and Adeleke et al. (2015) however follow (5) where covariate is time varying with introduction of a piece-wise exponential function. The latter generated baseline survival time for proportional and non-proportional by splitting the survival time into two non-overlapping domains $D_1(0,t_0]$ and $D_2[t_0, t)$. In using Weibull distribution, at *D*1 and *D*2, we obtained.

$$T = \left( \frac{-\log(U)}{\lambda \exp(\beta' x)} \right)^{\frac{1}{v}} \qquad (8)$$

$$T = \left( \frac{-\log(U) - \lambda \exp(\beta' x)t_0^v + \lambda \exp(\beta' x + \gamma)t_0^v}{\lambda \exp(\beta' x + \gamma)} \right)^{\frac{1}{v}} \qquad (9)$$

**Table 1**. Characterization of the Exponential, Weibull and Gompertz distributions, Bender et.al.(2005)

| Characteristic<br><br>Parameter | Exponential<br><br>Scale parameter $\lambda > 0$ | Weibull<br>Scale parameter $\lambda > 0$<br>Shape parameter $v > 0$ | Gompertz<br>Scale parameter $\lambda > 0$<br>Shape parameter $-\infty < \alpha < \infty$ |
|---|---|---|---|
| Hazard function | $h_o(t) = \lambda$ | $h_o(t) = \lambda v t^{v-1}$ | $h_o(\mathrm{t}) = \lambda \exp(\alpha)$ |
| Cumulative Hazard | $H_o(t) = \lambda t$ | $H_o(t) = \lambda t^v$ | $H_o(\mathrm{t}) = \frac{\lambda}{\alpha}\left(\exp(\alpha t) - 1\right)$ |
| Inverse Cum. Hazard | $H^{-1}{}_o(t) = \lambda^{-1} t$ | $H_o^{-1}(t) = (\lambda^{-1} t)^{\frac{1}{v}}$ | $H_o^{-1}(\mathrm{t}) = \frac{1}{\alpha}\log\left(\frac{\alpha}{\lambda} t + 1\right)$ |
| simulating survival time for time-invariant ($u$ ~$U$(0;1)) | $T = -\left(\frac{\log(\mathrm{u})}{\lambda \exp(\beta' x)}\right)$ | $T = -\left(\frac{\log(\mathrm{u})}{\lambda \exp(\beta' x)}\right)^{\frac{1}{v}}$ | $T = \frac{1}{\alpha}\log\left(1 - \frac{\alpha \log(\mathrm{u})}{\lambda \exp(\beta' x)}\right)$ |

The introduction of guarantee time into the baseline model, for Weibull distribution, given a density function with parameters $\lambda$ and v:

$$f_o(t; v, \lambda) \;= \lambda^v v(t - t_g)^{v-1} exp(-\lambda^v (t - t_g)^v) \qquad (10)$$

with Survival function:

$$S_o(t; v.\lambda) \;= exp(-\lambda^v (t - t_g)^v)$$

and hazard function:

$$h_o(t; v.\lambda) \;= v\lambda^v (t - t_g)^{v-1} \qquad (11)$$

Recall that the covariate is time varying in a piecewise form then, when $t < t_o : z(\mathrm{t}) = 0$ and when $t \geq t_o : z(\mathrm{t}) = 1$.Therefore for $t < t_o$ in $D_1$, and $t \geq t_0$ in $D_2$, we obtain the following:

$$\tau = \left(\frac{-\log(\mathrm{u}) + \lambda^v \exp(\beta'\mathrm{x})(-\mathrm{t}_g)^v}{\lambda^v \exp(\beta' x)}\right)^{\frac{1}{v}} \qquad (12)$$

$$\tau = \left(\frac{-\log(\mathrm{u}) + \lambda^v \exp(\beta'\mathrm{X})(-\mathrm{t}_g)^v - \lambda(1 - \exp(\gamma))\exp(\beta' X)(\mathrm{t}_o - \mathrm{t}_g)^v}{\lambda^v \exp(\beta'\mathrm{X} + \gamma)}\right)^{\frac{1}{v}} \qquad (13)$$

**Table 2.** Characterization of the Weibull distributions for both Bender et.al. (2005) and Adeleke et.al. (2015).

| **Characteristic** Parameter | **Weibull** Scale parameter $\lambda > 0$ Shape parameter $v > 0$ Bender et. al. | **Weibull** Scale parameter $\lambda > 0$ Shape parameter $v > 0$, $t_g \geq 0$ Proposed method |
|---|---|---|
| Hazard function | $h_o(t) = \lambda v t^{v-1}$ | $h_0(t) = v\lambda^v (t - t_g)^{v-1}$ |
| Cumulative Hazard | $H_o(t) = \lambda t^v$ | $H_0(t) = \lambda^v (t - t_g)^v$ |
| Inverse Cum. Hazard | $H_o^{-1}(t) = (\lambda^{-1} t)^{\frac{1}{v}}$ | $H_0(t)^{-1} = [\lambda^{-v}(t - t_g)]^{1/v}$ |
| simulating survival time for time-invariant (*u* ~*U*(0;1)) | $T = -\left(\dfrac{\log(u)}{\lambda \exp(\beta' x)}\right)^{\frac{1}{v}}$ | $\tau = \left(\dfrac{-\log(u) + \lambda^v \exp(\beta' x)(-t_g)^v}{\lambda^v \exp(\beta' x)}\right)^{\frac{1}{v}}, t < t_0$ |
| simulating survival time for time-varying (*u* ~*U*(0;1)) | | $\tau = \left(\dfrac{-\log(u) + \lambda^v \exp(\beta' X)(-t_g)^v - \lambda(1 - \exp(\gamma)) \exp(\beta' x)(t_o - t_g)^v}{\lambda^v \exp(\beta' X + \gamma)}\right)^{\frac{1}{v}}, t \geq t_0$ |

We now generate survival times that follow the model with guarantee time and a model without guarantee time (9) and (13) for (5) and (6) respectively via Weibull baseline hazard distribution (11). The simulation is in folds, the first is by generating censored data that follows Nonproportional hazards NPH model for three right censoring levels: Low (25%), Moderate (50%), High (80%). Four different samples size 10, 50, 100 and 500 were used to generate survival times in weeks. Given the censoring observations fixed at 42 weeks, the times were generated from a uniform distribution U(0,1). The time fixed covariate X was generated from N(0,1), the time-varying covariate Z(t) was generated as an interaction of Z with the function of time, where Z~B(n,0.5). With the introduction of guarantee time tg, the final survival times were generated using (8) through (13). We performed survival analysis on each generated data with 1,000 replications. We use the following true values of regression coefficients, = (2,-1) which are time fixed and time-varying parameters of the extended Cox models.

## Results and Discussion

The following are the results obtained from the simulation process following asymptotic properties of the estimators. Table 3 shows the results obtain using (5) and true parameter values. The estimated values slowly increase generally from Low (25%) to High (80%) levels of censoring. Deviation (bias) is high at small sample size and much lower at high sample size (N=500). Thus, comparing these estimates at different sample sizes, show that estimated values become more stable and centred around the true values. There are high values of bias at a moderate percentage of censoring (N=100). This could be due to sampling fluctuation as it does not appear to be so at other levels or sample sizes. Although, the general belief is that as sample size increases, our estimated parameters should explain better the true values. The values increase from low to high percentage of censoring slowly. Overestimation was noticed in most of the values, although is at its lowest in large samples.

**Table 3.** Mean values of the estimated coefficients of Extended Cox model.

| Sample sizes | % Cens | $\hat{\beta}$(2) | $\hat{\gamma}$(-1) |
|---|---|---|---|
| 10 | **LOW** | 3.718 | -1.081 |
| | **MODERATE** | 3.835 | -1.000 |
| | **HIGH** | 3.797 | -1.011 |
| 50 | **LOW** | 2.185 | -1.020 |
| | **MODERATE** | 2.177 | -1.019 |
| | **HIGH** | 2.159 | 1.022 |
| 100 | **LOW** | 1.893 | -1.009 |
| | **MODERATE** | 0.593 | -0.229 |
| | **HIGH** | 2.212 | -1.009 |
| 500 | **LOW** | 1.954 | -0.999 |
| | **MODERATE** | 1.931 | -1.002 |
| | **HIGH** | 0.195 | -0.004 |

The mean values of estimates obtained using (6), Table 4, showed that estimated values obtained were high at small sample sizes and low or close to true values as sample size increased. Also, sometimes the values increased slowly and sometimes it's stable and /or decreased slowly as the percentage of censoring increased. The degree of bias is very high compared to (5).

Thus, comparing these estimates at different levels of censoring and sample sizes, it can be seen that estimated values increased from low to high levels of censoring slowly (although appeared partially overestimated and underestimated).

Generally, in both tables, underestimation is more pronounced with the estimated parameters of the model (6); (Extended Cox model with guarantee time, see also Table 4). These characteristics of estimates are however used to observe the shifts.

**Table 4.** Mean values of the estimated coefficients of Extended-Cox model with guarantee time.

| Sample sizes | % Cens | $\hat{\beta}$(2) | $\hat{\gamma}$(-1) |
|---|---|---|---|
| 10 | **LOW** | 3.804 | -1.087 |
| | **MODERATE** | 3.923 | -1.005 |
| | **HIGH** | 3.884 | -1.016 |
| 50 | **LOW** | 2.235 | -1.025 |
| | **MODERATE** | 2.227 | -1.025 |
| | **HIGH** | 2.209 | 1.027 |
| 100 | **LOW** | 1.937 | -1.014 |
| | **MODERATE** | 0.606 | -0.230 |
| | **HIGH** | 2.263 | -1.014 |
| 500 | **LOW** | 2.0010 | -1.0020 |
| | **MODERATE** | 1.9990 | -0.9100 |
| | **HIGH** | 2.0010 | -1.0070 |

### *Performance Evaluation of Absolute Bias (AB)*

Table 5 shows the sensitivity of the models in relation to levels (percentage) of censoring. At low percentage of censoring, both models (with guarantee and without guarantee time) had minimum Absolute Bias (AB) except in one or two case(s) which may be due to sampling fluctuation. The reverse is the case when the percentage of censoring is moderate. This shows that both models are less biased especially at large sample (N=500). Also, at low and high levels of censoring, the estimated parameters of time fixed covariate in the model without guarantee time had minimum AB compared with models without guarantee time in a large sample. This was best achieved at a low percentage of censoring with minimal AB (0.0010). For time-varying covariate parameters, the model with guarantee time possessed the Least AB (0.0003) and was best achieved when the percentage of censoring was low, (see table 5).

Also, in terms of sample sizes, the least AB values existed when the sample size is 10 followed by 500 (large). The highest value of AB occurred at sample size 10 (moderate percentage of censoring) as indicated in table 5 and could be due to sampling fluctuation, as it does not appear to be so in other levels or sample sizes.

***Performance Evaluation of Mean Square Error (MSE) of the Estimated Coefficients***

Showing in the tables (6) and (7), below are the estimated mean square errors of the coefficients of models with and without guarantee time. Compared to the model without guarantee time (table 6), the model with guarantee time has a smaller MSE($\hat{\beta}$) for the parameter $\beta$ (table 6). However, the MSE($\beta$) from extended Cox model is greater than MSE($\beta$) from extended Cox model with guarantee time for all sample sizes except at 500. This is consistent at all levels of censoring. Meanwhile, for the MSE($\hat{\gamma}$) shown in the table (7), Extended cox model with guarantee time has smaller MSE ($\hat{\gamma}$) for the parameter $\hat{\gamma}$ at all sample sizes. Hence, this rules out extended Cox model in the presence of guarantee time.

Figure 1 is the charts showing the asymptotic behaviours of the model's parameters. From Figure1 (left), the behaviour of MSE($\hat{\beta}$), at Low, showed that MSE($\hat{\beta}$) decreased as sample size increased although the rate was faster between N = 10 and 50. At Moderate level, the estimates reduced first from N = 10 to 50 and fluctuated as sample size increased. Lastly, at high level, estimates reduced sharply from N = 10 to 100 where model with guarantee time behaved wildly and increased upward as sample size increased. This was actually unusual although initially, followed asymptotic property but faded out at N = 100. Hence, asymptotic properties held for models with guarantee time and without guarantee time at low and moderate levels of censoring but did not held at high level with guarantee time model.

**Table 5.** Absolute Biases of the estimated coefficients, N = 10-500.

| Sample sizes | | N=10 | | N=50 | | N=100 | | N=500 | |
|---|---|---|---|---|---|---|---|---|---|
| | | AB($\hat{\beta}$) | AB($\hat{\gamma}$) | AB($\hat{\beta}$) | AB($\hat{\gamma}$) | AB($\hat{\beta}$) | AB($\hat{\gamma}$) | AB($\hat{\beta}$) | AB($\hat{\gamma}$) |
| Ext Cox | **LOW** | 1.8036 | 0.0868 | 0.0635 | 0.0139 | 0.0637 | 0.0139 | 0.0010 | 0.0020 |
| | **MODERATE** | 1.9230 | 0.0047 | 1.3936 | 0.7697 | 1.395 | 0.7651 | 0.0010 | 1.9100 |
| | **HIGH** | 1.8843 | 0.0159 | 0.2631 | 0.0140 | 0.2647 | 0.0144 | 0.0010 | 0.0070 |
| Ext Cox wtGt | **LOW** | 1.7181 | 0.0003 | 0.1069 | 0.0089 | 0.1161 | 0.0080 | 0.0463 | 0.0007 |
| | **MODERATE** | 1.8348 | 0.0814 | 1.4072 | 0.7709 | 1.4112 | 0.7719 | 0.0687 | 0.0016 |
| | **HIGH** | 1.7970 | 0.0109 | 0.2123 | 0.0090 | 0.2203 | 0.0091 | 1.8047 | 0.9956 |

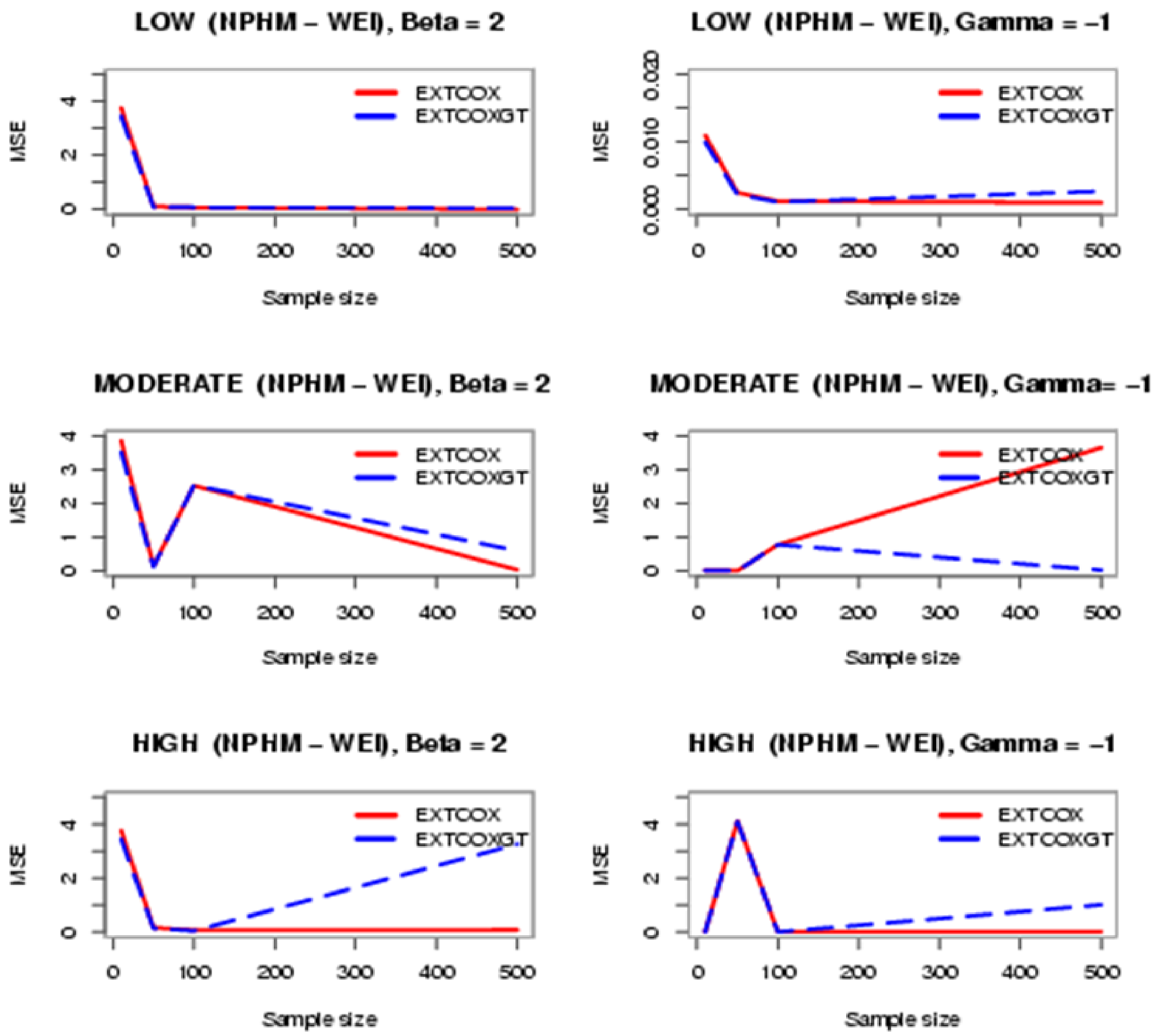


**Figure 1.** Asymptotic behaviour of MSE ( $\hat{\beta}$ ),and MSE( $\hat{\gamma}$ ).

**Table 6.** Summary of Mean Square Error (MSE( $\hat{\beta}$ =2)) of The Estimated Coefficients of models

| **Level** | **Low MSE ( $\hat{\beta}$ )** | | | | **Moderate MSE ( $\hat{\beta}$ )** | | | | **High MSE ( $\hat{\beta}$ )** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample Size | 10 | 50 | 100 | 500 | 10 | 50 | 100 | 500 | 10 | 50 | 100 | 500 |
| Ext | 3.7666 | 0.1110 | 0.0769 | 0.0010 | 3.8541 | 0.1473 | 2.5213 | 0.0330 | 3.7696 | 0.1613 | 0.0700 | 0.0790 |
| Ext wtgt | 3.4426 | 0.0874 | 0.0811 | 0.0516 | 3.5157 | 0.1228 | 2.5337 | 0.5944 | 3.4384 | 0.1379 | 0.0458 | 3.2584 |

**Table 7.** Summary of Mean Square Error (MSE( $\hat{\gamma}$ = -1)) of The Estimated Coefficients of models.

| **Level** | **Low MSE( $\hat{\gamma}$ )** | | | | **Moderate MSE( $\hat{\gamma}$ )** | | | | **High MSE( $\hat{\gamma}$ )** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample Size | 10 | 50 | 100 | 500 | 10 | 50 | 100 | 500 | 10 | 50 | 100 | 500 |
| Ext | 0.0110 | 0.0025 | 0.00122 | 0.0010 | 0.0098 | 0.0025 | 0.7753 | 3.6499 | 0.0103 | 4.1098 | 0.0011 | 0.0119 |
| Ext wtgt | 0.0100 | 0.0022 | 0.00109 | 0.0027 | 0.0097 | 0.0023 | 0.7753 | 0.0213 | 0.0101 | 4.0892 | 0.0010 | 1.0099 |

Also, in Figure 1 (right), the behaviour was quite interesting, MSE($\hat{\gamma}$) of model with guarantee time (High) increased as sample size increased, this was as opposed to the asymptotic property, while model without guarantee time satisfied the asymptotic property at Low and High levels. The result was different at moderate level where model with guarantee time satisfied the asymptotic property. The reverse is the case with model without guarantee time. Hence, the asymptotic property held for extended Cox only at Low and High levels of censoring and model with guarantee time only at low and moderate.

***Behaviour of Mean Square Error at Different Levels of Censoring***

Figure 2 and Figure 3 show the behaviour of MSE of time fixed and time varying coefficient ($\hat{\beta}$, $\hat{\gamma}$) at different levels of censoring. At sample size 10, MSE's ($\hat{\beta}$) are high for both models and as percentage of censoring increases, it also increases. While in the case of MSE's ($\hat{\gamma}$), tends to zero.

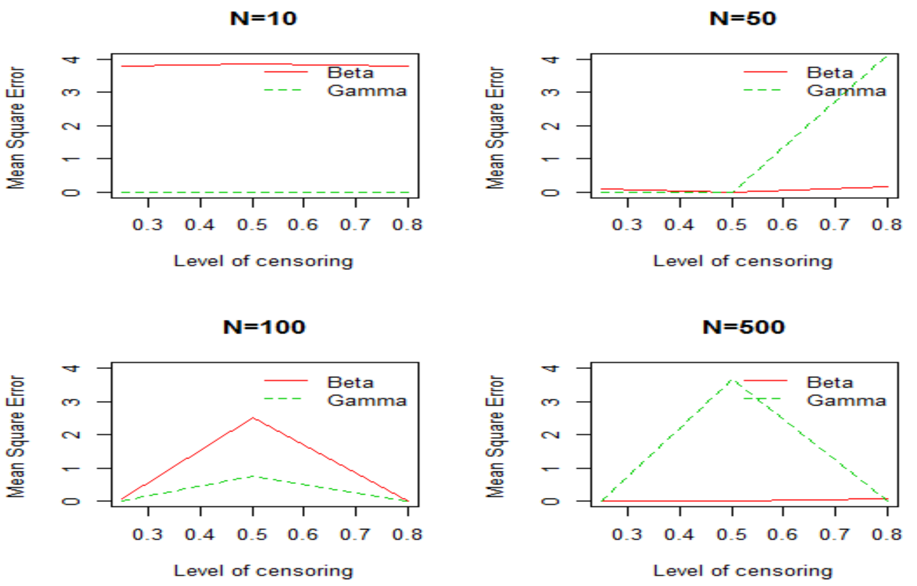


**Figure 2.** Behaviour of MSE from Extended Cox model at different levels of censoring.

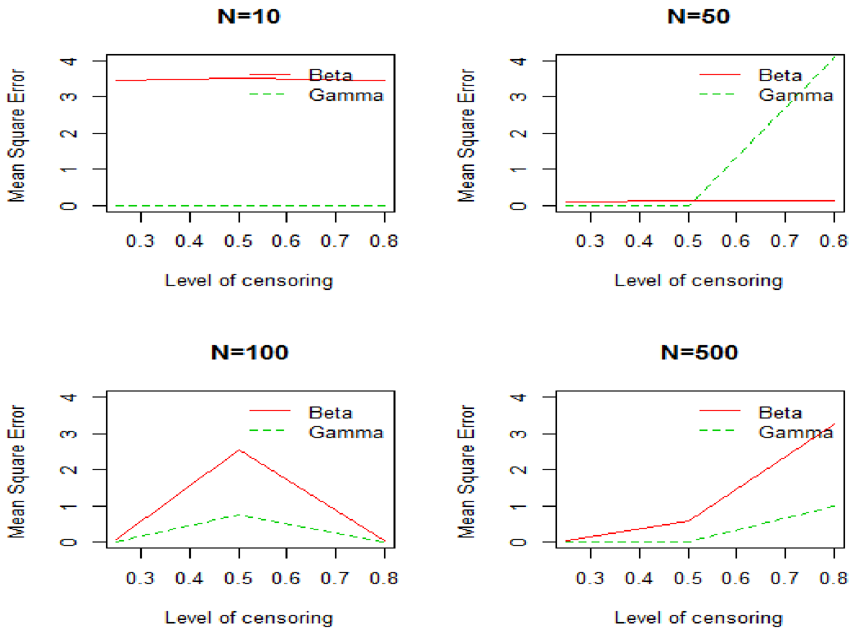


**Figure 3.** Behaviour of MSE from Extended Cox model with guarantee time at different levels of censoring

The case is different at higher sample sizes (N=500) where MSE's increases as the percentage of censoring increases (model with guarantee time), see Figures 2 and 3 above. The implication is that at low levels of censoring, both models performed better and have minimum MSE and at high levels, it is fair with a model without guarantee time (Extended Cox model).

### *Application*

We present an example using real-life data on pregnancy duration obtained from a retrospective study conducted in Nigeria between January to August, 2015 at Adeoyo Maternity Teaching Hospital in Nigeria. In this study, records containing the history of pregnant women registered and monitored till delivery were randomly selected. From the record, the following observations were taken: Date of Admission (DOA), Date of Delivery (DOD). Estimated date of Delivery (EDD), Last menstrual period (LMP), Scan dates, Presence or absence of Meconium stained (ms=1, present and ms=0, absent) and blood pressure (BP = 1 if High, 0 if normal and -1 if low). By considering the report of WHO (2012) as regards the moderate and late preterms which occurs from 32to 37 weeks of gestational age, we, however, choose our guarantee time as 32. This does not outrightly ignore the births or deliveries before 32 weeks (i.e guarantee time) but trying to look for a safe time or minimum time within which no birth can occur (i.e saving pregnancy women from having abortion, complications, early preterm births). Preterm infants are at greater risk than term infants, for complications of mortality, prematurity, feeding, hypoglycaemia, jaundice, temperature instability, respiratory distress, Oddie et al. (2005), Engle et al. (2007) and William et.al. (2008). Note: the knot used during analysis under the step function is $t_0$= 32. i.e:

$$g(t) = \begin{cases} 1, & t \geq 32 \\ 0 & t < 32 \end{cases}$$

We observed that a high Estimates of Extended Cox model with and without guarantee time obtained from data on pregnancy duration were presented in table 3. blood pressure pregnant women had relatively lower hazard ratio of 0:929 (se($\hat{\beta}$) = 0:090), (P = 0:410) although its effect was not significant. Her estimated hazard ratio for error in dating (Err) when pregnancy period was equal to or exceeded 32 weeks was 0:652(se( $\hat{\beta}$ ) = 0:087), (P <0:0001). This was approximately 35% decreased in hazard for the effect of pregnant women with lower error in dating per week increment in time to delivery, especially at age less than 32 weeks. That is, there is about 891% increase in hazard ratio for the effect of pregnant women with higher error in dating per week increment in time to delivery especially at age less than 32 weeks. The presence of meconium stained during the pregnancy indicated a highly significant (P <0:0001) hazard ratio of 0:266(Se( $\hat{\beta}$ ) = 0:122) when time to delivery exceeded 32 weeks. In the same vein, a highly significant (P <0:0001) hazard ratio of 74:000(Se( $\hat{\beta}$ ) = 0:495, see Table 8) effect when time to delivery was less than 32 weeks.

Estimates based on Extended Cox model with guarantee time (6). The presence of high blood pressure with hazard ratio of 0:998(Se( $\hat{\beta}$ ) = 0:089) had no significant effect (P = 0:984) on time to delivery. Estimates of error in dating showed a non-significant (P = 0:089) hazard ratio of 1:157(Se( $\hat{\beta}$ ) = 0:086) when event time equal to or exceeded 32 weeks, in contrast to a highly significant (P <0:001) hazard ratio of 0:273(Se( $\hat{\beta}$ ) = 0:257) when was less than 32 weeks.

**Table 8**. Estimates from Extended Cox (Non-proportional hazard) Models.

| **Variables** | **Without Guarantee Time** | | | **With Guarantee Time** | | |
|---|---|---|---|---|---|---|
| | Exp( $\hat{\beta}$ ) | Se( $\hat{\beta}$ ) | P | Exp( $\hat{\beta}$ ) | Se( $\hat{\beta}$ ) | P |
| High BP | 0.929 | 0.090 | 0.410 | 0.998 | 0.089 | 0.984 |
| Error in Dating(gt) | 0.625 | 0.087 | <0:0001 | 1.157 | 0.086 | 0.089 |
| Error in Dating($gt_1$) | 8.910 | 0.476 | <0:0001 | 0.273 | 0.257 | <0:0001 |
| Meconium Stained (gt) | 0.266 | 0.122 | <0:0001 | 0.522 | 0.104 | <0:0001 |
| Meconium Stained ($gt_1$) | 74.000 | 0.495 | <0:0001 | 1.007 | 0.507 | 0.989 |
| Global | <0.0001 | | | 0.266 | | |
| AIC (5df) | 7086.983 | | | 7084.042 | | |
| -2LOGLIK | 7076.983 | | | 7074.04 | | |

Also the presence of Meconium stained showed a significant (P <0:001) hazard ratio of 0:522(Se( ) = 0:104) effect when time exceeded 32 weeks while in contrast, a non-significant (P = 0:988) hazard ratio of 1:007(Se( $\hat{\beta}$ ) = 0:507) when time was less than 32 weeks, Table 8. In the same vein, comparison of the estimates with and without guarantee time showed that model with guarantee time is preferable the more considering model selection criterion. The log likelihood ratio (7074:04) less (7076:983) favoured a model with guarantee time, and AIC = 7084:042 obtained from a model with guarantee time was preferred to model without guarantee time with AIC = 7086:983. Hence, model with guarantee time performed better.

## Conclusion

Generally, the framework developed here is for modelling the nature of potentially semiparametric non-proportional models with and without guarantee time and the usefulness of the methods developed in our work go well beyond the scope of simulations but extends to wide range applications. Coupled with the work in Bender et. al. (2005) and Austin (2012), our work extends inference methods to a wide range of statistics. More so, further research may focus on inferences problems and on joint testing and modelling for multiple covariates, time-varying covariate along the lines. In summary, the introduction of guarantee time is a new concept and covariate effects in non-proportional hazard models pose big problems that need lots of attention. The models developed are useful in many applications. However, our work also highlighted the need for further research in other directions.

## Acknowledgement

I am using this medium to thank all that contributed to this work in one way or the other. Special thanks to Prof. Benjamin A. Oyejola (Department of Statistics, University of Ilorin, Ilorin, Kwara State, Nigeria), for his erudite suggestions an advice as well as Dr.Akindele (Department of Psychology, Olabisi Onabanjo University, Ago-Iwoye, Nigeria) for his support during data collection.Thank you all for helping me.

## References

Adeleke, K. A., Abiodun A. A. & Ipinyomi, R.A. (2015). Semi-Parametric Non-Proportional Hazard Model With Time Varying Covariate. *Journal of Modern Applied Statistical Methods, 14*(2), 68 – 87.

Arasan, J. & Lunn, M. (2009). Survival model of a parallel system with dependent failures and time varying covariates. *Journal of Pharmaceutical and Scientific Innovation, 139*, 944-951.

Ata, N. & Sozer, T. M. (2007). Cox regression models cancer survival data. *Hacettepe Journal of Mathematics and Statistics, 36*(2), 157-167.

Austin, P.C. (2012). Generating survival times to simulate Cox proportional hazards models with time-varying covariates. *Wiley Statist.Med., 31*,157-167.

Bender, R., Augustin, T. & Blettner, M. (2005). Generating survival times to simulate cox proportional hazards models. *Statistics in Medicine, 25*(11), 1713-1723. doi:10.1002/sim.2369.

Engle, W. A., Tomashek, K. M. & Wallman, C. (2007). "Late-preterm" infants: a population at risk. *Pediatrics, 120*(6), 1390-1401.

Engle, W. A., & Kominiarek, M. A. (2008). Late preterm infants, early term infants, and timing of elective deliveries. *Clinics in perinatology, 35*(2), 325-341.

Kronborg, D. & Aaby, P. (1990). Piecewise Comparison of Survival Function in Stratied Proportional Hazard Models. *Biometrics, 46*, 375-380.

Lee, E.T. & Wang, J.W. (2003). Statistical Methods for survival Data Analysis. John Wiley, New Jersey.

Leemis, L.M., Shihand, L.H. & Reynertson, K. (1990). Variate generation for accelerated life and proportional hazards models with time dependent covariates. *Statistics and Probability Letters, 10*(6), 335-339. doi:10.1016/0167-7152(90)90052-9.

Leemis, L.M., (1987). Variate generation for accelerated life and proportional hazards models. *Operations Research, 35*(6): 892-894. doi:10.1287/opre.35.6.892.

Mood A.M., Graybill F.A. & Boes D.C. (1974). Introduction to the Theory of Statistics. McGraw-Hill: New York.

Oddie, S. J., Hammal, D., Richmond, S. & Parker, L. (2005). Early discharge and readmission to hospital in the first month of life in the Northern Region of the UK during 1998: a case cohort study. *Archives of disease in childhood, 90*(2), 119-124.

Robert, C. P. & Casella, G. (2009). Introducing Monte Carlo Methodswith R, Springer-Verlag.

Scheike, T. (2004). Time varying effect in survival analysis. *Handbook of statistics Elsevier, 23*, 61-85. doi:10.1016/S0169-7161(03)23004-2

Shihand, L.H. & Leemis, L. M. (1993). Variate generation for a nonhomogenous Poisson process with time dependent covariates. *Journal of Statistical Computation and Simulation, 44*, 165-186. doi:10.1080/00949659308811457.

Sylvester, M.P. & Abrahamowicz, M. (2007). Comparison of algorithms to generate event times conditional on time-dependent covariates. *Statistics in Medicine, 27*(14), 2618-2634. doi:10.1002.

Wong, H., Ip, W. & Zhang, R.Q. (2006). Varying-coefcient single-index model. Csda5214581476.

Zhou, M. (2001). Understanding the cox regression models with time change covariates. *The American Statistician, 55*(2),153-155.

**How to cite this paper:**
Adeleke, K. A., Abiodun, A. A. & Ipinyomi, R. A. (2018). Extended Cox Modelling of Survival Data with Guarantee Time. *Malaysian Journal of Applied Sciences, 3*(2), 21-33.